

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: BONNEFOY, JEAN-YVES
GAUCHAT, JEAN-FRANCOIS
- (ii) TITLE OF INVENTION: SUBSTANCES AND THEIR USES
- (iii) NUMBER OF SEQUENCES: 9
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: NIXON & VANDERHYE P.C.
 - (B) STREET: 1100 NORTH GLEBE ROAD
 - (C) CITY: ARLINGTON
 - (D) STATE: VIRGINIA
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 22201-4714
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: MS Word
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE: 29-SEP-2003
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 09/545,002
 - (B) FILING DATE: 06-APR-2000
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/969,125
 - (B) FILING DATE: 12-NOV-1997
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: GB 9625899.1
 - (B) FILING DATE: 13-DEC-1996

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (v) FRAGMENT TYPE: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Tyr Xaa Xaa Gln
1

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Trp Ser Xaa Trp Ser
1 5

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(v) FRAGMENT TYPE: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTGAGCTACA TGAAGTGTTC TTGGCTCCCT

30

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(v) FRAGMENT TYPE: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CAGAGTTTGT CATCCTCATA GCATAACTTA

30

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(v) FRAGMENT TYPE: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATACCACTC CCGACACTAA CTATACTCTC

30

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(v) FRAGMENT TYPE: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGAGAAATGG CTTCGTTTG CTGGCTATC

30

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(v) FRAGMENT TYPE: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TACCATGTCT CTGTATATGG AAAGTCTTCA

30

(2) INFORMATION FOR SEQ ID NO:8

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4038 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(v) FRAGMENT TYPE: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 43..1323

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8

GCCAAGGCTC CAGCCCGGCC GGGCTCCGAG GCGAGAGGCT GC ATG GAG TGG CCG
Met Glu Trp Pro
1

54

GCG CGG CTC TGC GGG CTG TGG GCG CTG CTG CTC TGC GCC GGC GGC GGG
Ala Arg Leu Cys Gly Leu Trp Ala Leu Leu Cys Ala Gly Gly Gly
5 10 15 20

102

GGC GGG GGC GGG GGC GCC GCG CCT ACG GAA ACT CAG CCA CCT GTG ACA
Gly Gly Gly Gly Gly Ala Ala Pro Thr Glu Thr Gln Pro Pro Val Thr
25 30 35

150

AAT TTG AGT GTC TCT GTT GAA AAC CTC TGC ACA GTA ATA TGG ACA TGG
Asn Leu Ser Val Ser Val Glu Asn Leu Cys Thr Val Ile Trp Thr Trp
40 45 50

198

AAT CCA CCC GAG GGA GCC AGC TCA AAT TGT AGT CTA TGG TAT TTT AGT

246

Asn Pro Pro Glu Gly Ala Ser Ser Asn Cys Ser Leu Trp Tyr Phe Ser	
55 60 65	
CAT TTT GGC GAC AAA CAA GAT AAG AAA ATA GCT CCG GAA ACT CGT CGT	294
His Phe Gly Asp Lys Gln Asp Lys Lys Ile Ala Pro Glu Thr Arg Arg	
70 75 80	
TCA ATA GAA GTA CCC CTG AAT GAG AGG ATT TGT CTG CAA GTG GGG TCC	342
Ser Ile Glu Val Pro Leu Asn Glu Arg Ile Cys Leu Gln Val Gly Ser	
85 90 95 100	
CAG TGT AGC ACC AAT GAG AGT GAG AAG CCT AGC ATT TTG GTT GAA AAA	390
Gln Cys Ser Thr Asn Glu Ser Glu Lys Pro Ser Ile Leu Val Glu Lys	
105 110 115	
TGC ATC TCA CCC CCA GAA GGT GAT CCT GAG TCT GCT GTG ACT GAG CTT	438
Cys Ile Ser Pro Glu Gly Asp Pro Glu Ser Ala Val Thr Glu Leu	
120 125 130	
CAA TGC ATT TGG CAC AAC CTG AGC TAC ATG AAG TGT TCT TGG CTC CCT	486
Gln Cys Ile Trp His Asn Leu Ser Tyr Met Lys Cys Ser Trp Leu Pro	
135 140 145	
GGA AGG AAT ACC AGT CCC GAC ACT AAC TAT ACT CTC TAC TAT TGG CAC	534
Gly Arg Asn Thr Ser Pro Asp Thr Asn Tyr Thr Leu Tyr Tyr Trp His	
150 155 160	
AGA AGC CTG GAA AAA ATT CAT CAA TGT GAA AAC ATC TTT AGA GAA GGC	582
Arg Ser Leu Glu Lys Ile His Gln Cys Glu Asn Ile Phe Arg Glu Gly	
165 170 175 180	
CAA TAC TTT GGT TGT TCC TTT GAT CTG ACC AAA GTG AAG GAT TCC AGT	630
Gln Tyr Phe Gly Cys Ser Phe Asp Leu Thr Lys Val Lys Asp Ser Ser	
185 190 195	
TTT GAA CAA CAC AGT GTC CAA ATA ATG GTC AAG GAT AAT GCA GGA AAA	678
Phe Glu Gln His Ser Val Gln Ile Met Val Lys Asp Asn Ala Gly Lys	
200 205 210	
ATT AAA CCA TCC TTC AAT ATA GTG CCT TTA ACT TCC CGT GTG AAA CCT	726
Ile Lys Pro Ser Phe Asn Ile Val Pro Leu Thr Ser Arg Val Lys Pro	
215 220 225	
GAT CCT CCA CAT ATT AAA AAC CTC TCC TTC CAC AAT GAT GAC CTA TAT	774
Asp Pro Pro His Ile Lys Asn Leu Ser Phe His Asn Asp Asp Leu Tyr	
230 235 240	
GTG CAA TGG GAG AAT CCA CAG AAT TTT ATT AGC AGA TGC CTA TTT TAT	822
Val Gln Trp Glu Asn Pro Gln Asn Phe Ile Ser Arg Cys Leu Phe Tyr	
245 250 255 260	
GAA GTA GAA GTC AAT AAC AGC CAA ACT GAG ACA CAT AAT GTT TTC TAC	870
Glu Val Glu Val Asn Asn Ser Gln Thr Glu Thr His Asn Val Phe Tyr	
265 270 275	
GTC CAA GAG GCT AAA TGT GAG AAT CCA GAA TTT GAG AGA AAT GTG GAG	918
Val Gln Glu Ala Lys Cys Glu Asn Pro Glu Phe Glu Arg Asn Val Glu	
280 285 290	
AAT ACA TCT TGT TTC ATG GTC CCT GGT GTT CTT CCT GAT ACT TTG AAC	966
Asn Thr Ser Cys Phe Met Val Pro Gly Val Leu Pro Asp Thr Leu Asn	
295 300 305	
ACA GTC AGA ATA AGA GTC AAA ACA AAT AAG TTA TGC TAT GAG GAT GAC	1014

Thr	Val	Arg	Ile	Arg	Val	Lys	Thr	Asn	Lys	Leu	Cys	Tyr	Glu	Asp	Asp	
310						315				320						
AAA	CTC	TGG	AGT	AAT	TGG	AGC	CAA	GAA	ATG	AGT	ATA	GGT	AAG	AAG	CGC	1062
Lys	Leu	Trp	Ser	Asn	Trp	Ser	Gln	Glu	Met	Ser	Ile	Gly	Lys	Lys	Arg	
325					330					335					340	
AAT	TCC	ACA	CTC	TAC	ATA	ACC	ATG	TTA	CTC	ATT	GTT	CCA	GTC	ATC	GTC	1110
Asn	Ser	Thr	Leu	Tyr	Ile	Thr	Met	Leu	Leu	Ile	Val	Pro	Val	Ile	Val	
				345					350					355		
GCA	GGT	GCA	ATC	ATA	GTA	CTC	CTG	CTT	TAC	CTA	AAA	AGG	CTC	AAG	ATT	1158
Ala	Gly	Ala	Ile	Ile	Val	Leu	Leu	Leu	Tyr	Leu	Lys	Arg	Leu	Lys	Ile	
			360				365						370			
ATT	ATA	TTC	CCT	CCA	ATT	CCT	GAT	CCT	GGC	AAG	ATT	TTT	AAA	GAA	ATG	1206
Ile	Ile	Phe	Pro	Pro	Ile	Pro	Asp	Pro	Gly	Lys	Ile	Phe	Lys	Glu	Met	
		375					380					385				
TTT	GGA	GAC	CAG	AAT	GAT	GAT	ACT	CTG	CAC	TGG	AAG	AAG	TAC	GAC	ATC	1254
Phe	Gly	Asp	Gln	Asn	Asp	Asp	Thr	Leu	His	Trp	Lys	Lys	Tyr	Asp	Ile	
		390				395					400					
TAT	GAG	AAG	CAA	ACC	AAG	GAG	GAA	ACC	GAC	TCT	GTA	GTG	CTG	ATA	GAA	1302
Tyr	Glu	Lys	Gln	Thr	Lys	Glu	Glu	Thr	Asp	Ser	Val	Val	Leu	Ile	Glu	
405					410				415					420		
AAC	CTG	AAG	AAA	GCC	TCT	CAG	TGATGGAGAT	AATTTATTTT	TACCTTCACT							1353
Asn	Leu	Lys	Lys	Ala	Ser	Gln										
				425												
GTGACCTTGA	GAAGATTCTT	CCCATTCTCC	ATTGTGTATC	TGGGAACTTA	TTAAATGGAA											1413
ACTGAAACTA	CTGCACCATT	TAAAAACAGG	CAGCTCATAA	GAGCCACAGG	TCTTTATGTT											1473
GAGTCGCGCA	CCGAAAAACT	AAAAATAATG	GGCGCTTTGG	AGAAGAGTGT	GGAGTCATTC											1533
TCATTGAATT	ATAAAAAGCCA	GCAGGCTTCA	AACTAGGGGA	CAAAGCAAAA	AGTGATGATA											1593
GTGGTGGAGT	TAATCTTATC	AAGAGTTGTG	ACAACCTCCT	GAGGGATCTA	TACTTGTCTT											1653
GTGTTCTTGT	TGTCAACATG	AACAAATTTT	ATTGTGAGGG	GAACCTATTT	GGGGTGCAAA											1713
TGCTAATGTC	AAACTTGTAGT	CACAAAGAAC	ATGTAGAAAA	CAAAATGGAT	AAAATCTGAT											1773
ATGTATTGTT	TGGGATCCCTA	TTGAACCATG	TTTGTGGCTA	TTAAACTCT	TTTAACAGTC											1833
TGGGCTGGGT	CCGGTGGCTC	ACGCCTGTAA	TCCCAGCAAT	TTGGGAGTCC	GAGGCGGGCG											1893
GATCACTCGA	GGTCAGGAGT	TCCAGACCAG	CCTGACCAAA	ATGGTGAAAC	CTCCTCTCTA											1953
CTAAAACTAC	AAAAATTAC	TGGGTGTGGT	GGCGCGTGCC	TGTAATCCCA	GCTACTCGGG											2013
AAGCTGAGGC	AGGTGAATTG	TTTGAACCTG	GGAGGTGGAG	GTTCAGTGA	GCAGAGATCA											2073
CACCACTGCA	CTCTAGCCTG	GGTGACAGAG	CAAGACTCTG	TCTAAAAAAC	AAAAACAAAC											2133
AAAAACAAAC	AAAAAACCT	CTTAATATTC	TGGAGTCATC	ATTCCCTTCG	ACAGCATTTT											2193
CCTCTGCTTT	GAAGCCCCA	GAAATCAGTG	TTGGCCATGA	TGACAACTAC	AGAAAAACCA											2253
GAGGCGAGCTT	CTTTGCCAAG	ACCTTTCAAA	GCCATTTTAG	GCTGTTAGGG	CGAGTGGAGG											2313
TAGAATGACT	CCTTGGGTAT	TAGAGTTTCA	ACCATGAAGT	CTCTAACAT	GTATTTTCTT											2373

CACCTCTGCT ACTCAAGTAG CATTACTGT GTCTTTGGTT TGTGCTAGGC CCCC GG GTGT	2433
GAAGCAGACA CCCCTCCAG GGGTTACAG TCTATTTGAG ACTCCTCAGT TCTTGCCACT	2493
TTTTTTTTTA ATCTCCACCA GTCATTTTTC AGACCTTTTA ACTCCTCAAT TCCAACACTG	2553
ATTTCCCTTT TTGCATTCTC CCTCCTTCCC TTCCTGTAG CCTTTTGACT TTCATTGGAA	2613
ATTAGGATGT AAATCTGCTC AGGAGACCTG GAGGAGCAGA GGATAATTAG CATCTCAGGT	2673
TAAGTGTGAG TAATCTGAGA AACAATGACT AATCTTGCA TATTTGTAA CTTCATGTG	2733
AGGGTTTTTA GCATTGATAT TTGTGCATT TCTAAACAGA GATGAGGTGG TATCTCAGC	2793
TAGAACATTG GTATTCGCTT GAGAAAAAAA GAATAGTTGA ACCTATTCTT CTTCTTTTAC	2853
AAGATGGGTC CAGGATTCTT CTTTTCTCTG CCATAAATGA TTAATTAAT AGCTTTTGTG	2913
TCTTACATTG GTAGCCAGCC AGCCAAGGCT CTGTTTATGC TTTTGGGGGG CATATATTGG	2973
GTTCCATTCT CACCTATCCA CACAACATAT CCGTATATAT CCCCTCTACT CTTACTTCCC	3033
CCAAATTTAA AGAAGTATGG GAAATGAGAG GCATTTCCCC CACCCCATTT CTCTCTCAC	3093
ACACAGACTC ATATTACTGG TAGGAACCTG AGAACTTTAT TTCCAAGTTG TTCAAACATT	3153
TACCAATCAT ATTAATACAA TGATGCTATT TGCAATTCCT GCTCCTAGGG GAGGGGAGAT	3213
AAGAAACCT CACTCTCTAC AGGTTTGGGT ACAAGTGGCA ACCTGCTTCC ATGGCCGTGT	3273
AGAAGCATGG TGCCCTGGCT TCTCTGAGGA AGCTGGGGTT CATGACAATG GCAGATGTAA	3333
AGTTATTCTT GAAGTCAGAT TGAGGCTGGG AGACAGCCGT AGTAGATGT CTACTTTGTT	3393
CTGCTGTCTT CTAGAAAGAA TATTTGTTT TCCTGTATAG GAATGAGATT AATCCITTTC	3453
CAGGTATTTT ATAATTCTGG GAAGCAAAAC CCATGCCTCC CCCTAGCCAT TTTTACTGTT	3513
ATCCTATTTA GATGGCCATG AAGAGGATGC TGTGAAATTC CCAACAAACA TTGATGCTGA	3573
CAGTCATGCA GTCTGGGAGT GGGGAAGTGA TCTTTTGTTT CCATCCTCTT CTTTTCAG	3633
TAAAATAGCT GAGGGAAAAA GGAGGGAAAA GGAAGTTATG GGAATACCTG TGGTGTTGT	3693
GATCCCTAGG TCTTGGGAGC TCTTGAGGT GTCTGTATCA GTGGATTTC CATCCCCTGT	3753
GGGAAATTAG TAGGCTCAT TACTGTTTTA GGTCTAGCCT ATGTGGATT TTTCTTAACA	3813
TACCTAAGCA AAGCCAGTGT CAGGATGGTA ATCTTATTC TTTCGTTTCA TTAAGTTTTT	3873
CCCTTCATCT GGGCACTGAA GGGATATGTG AAACAATGTT AACATTTTGT GTAGCTTCA	3933
ACCAGGGATT GTTCTGTGTT AACTTCTTAT AGGAAAGCTT GAGTAAATA AATATTGTCT	3993
TTTTGTATGT CAAGCGGGCC GCCACCGCGG TGGAAACTCC AGCTT	4038

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 427 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

Met Glu Trp Pro Ala Arg Leu Cys Gly Leu Trp Ala Leu Leu Leu Cys
 1             5             10             15
Ala Gly Gly Gly Gly Gly Gly Gly Gly Ala Ala Pro Thr Glu Thr Gln
          20             25             30
Pro Pro Val Thr Asn Leu Ser Val Ser Val Glu Asn Leu Cys Thr Val
          35             40             45
Ile Trp Thr Trp Asn Pro Pro Glu Gly Ala Ser Ser Asn Cys Ser Leu
          50             55             60
Trp Tyr Phe Ser His Phe Gly Asp Lys Gln Asp Lys Lys Ile Ala Pro
 65             70             75             80
Glu Thr Arg Arg Ser Ile Glu Val Pro Leu Asn Glu Arg Ile Cys Leu
          85             90             95
Gln Val Gly Ser Gln Cys Ser Thr Asn Glu Ser Glu Lys Pro Ser Ile
          100            105            110
Leu Val Glu Lys Cys Ile Ser Pro Glu Gly Asp Pro Glu Ser Ala
          115            120            125
Val Thr Glu Leu Gln Cys Ile Trp His Asn Leu Ser Tyr Met Lys Cys
          130            135            140
Ser Trp Leu Pro Gly Arg Asn Thr Ser Pro Asp Thr Asn Tyr Thr Leu
          145            150            155            160
Tyr Tyr Trp His Arg Ser Leu Glu Lys Ile His Gln Cys Glu Asn Ile
          165            170            175
Phe Arg Glu Gly Gln Tyr Phe Gly Cys Ser Phe Asp Leu Thr Lys Val
          180            185            190
Lys Asp Ser Ser Phe Glu Gln His Ser Val Gln Ile Met Val Lys Asp
          195            200            205
Asn Ala Gly Lys Ile Lys Pro Ser Phe Asn Ile Val Pro Leu Thr Ser
          210            215            220
Arg Val Lys Pro Asp Pro Pro His Ile Lys Asn Leu Ser Phe His Asn
          225            230            235            240
Asp Asp Leu Tyr Val Gln Trp Glu Asn Pro Gln Asn Phe Ile Ser Arg
          245            250            255
Cys Leu Phe Tyr Glu Val Glu Val Asn Asn Ser Gln Thr Glu Thr His
          260            265            270
Asn Val Phe Tyr Val Gln Glu Ala Lys Cys Glu Asn Pro Glu Phe Glu
          275            280            285
Arg Asn Val Glu Asn Thr Ser Cys Phe Met Val Pro Gly Val Leu Pro
          290            295            300
Asp Thr Leu Asn Thr Val Arg Ile Arg Val Lys Thr Asn Lys Leu Cys
          305            310            315            320

```

Tyr Glu Asp Asp Lys Leu Trp Ser Asn Trp Ser Gln Glu Met Ser Ile
 325 330 335
 Gly Lys Lys Arg Asn Ser Thr Leu Tyr Ile Thr Met Leu Leu Ile Val
 340 345 350
 Pro Val Ile Val Ala Gly Ala Ile Ile Val Leu Leu Leu Tyr Leu Lys
 355 360 365
 Arg Leu Lys Ile Ile Ile Phe Pro Pro Ile Pro Asp Pro Gly Lys Ile
 370 375 380
 Phe Lys Glu Met Phe Gly Asp Gln Asn Asp Asp Thr Leu His Trp Lys
 385 390 395 400
 Lys Tyr Asp Ile Tyr Glu Lys Gln Thr Lys Glu Glu Thr Asp Ser Val
 405 410 415
 Val Leu Ile Glu Asn Leu Lys Lys Ala Ser Gln
 420 425